

O tópico estilo-chinês e o *status* tipológico do PB: um estudo de rastreamento ocular

Letícia Gonçalves¹

Lorrane Medeiros²

RESUMO:

Com o objetivo de oferecer evidências que sustentem a hipótese do português do Brasil (PB) como língua mista, isto é, uma língua que, além de apresentar proeminência de sujeito, também apresenta proeminência de tópico, o presente estudo analisa o processamento de tópicos estilo-chinês por falantes nativos de PB, através de rastreamento ocular. Tal experimento apresenta *design* 4, onde o fator “Tipo de referente”, com 4 níveis, gerou as seguintes condições experimentais:

- a) Aquelas florestas, os arbustos são grandes, por isso eu gostei delas. (TOP)
- b) Aquelas florestas, os arbustos são grandes, por isso eu gostei deles. (SUJ)
- c) Aquelas florestas, as árvores são grandes, por isso eu gostei delas. (AMB)
- d) Aquelas florestas, os arbustos são grandes, por isso eu gostei _____. (NUL)

Conduziu-se um teste de leitura, onde a hipótese era a de que, numa língua de tópico, o que controlaria a correferência numa estrutura de tópico-comentário seria o elemento tópico e não o sujeito (cf. Li & Thompson, 1976). Logo, espera-se que os participantes considerem que o pronome anafórico seja correferente ao tópico. Os tempos de resposta às questões interpretativas indicaram um maior tempo de processamento nas condições “nulo” e “ambíguo”. Os resultados indicam, ainda, que, quando não há uma relação explícita entre o elemento correferente e o pronome, 65% das escolhas se dá pelo tópico, e não pelo sujeito. Tais dados corroboram achados de Medeiros (2021) e refutam Kenedy & Mota (2012), que avaliaram o PB como língua de proeminência de sujeito, quando investigaram o mesmo fenômeno.

PALAVRAS-CHAVE: Psicolinguística. Tópico-comentário. Rastreamento ocular. PB.

¹ Licencianda em Letras: Português-Inglês pela Universidade Federal do Rio de Janeiro (UFRJ). E-mail: leticiabrigoncalves@gmail.com ORCID: 0009-0006-3674-5567.

² Doutora em Linguística pelo Programa de Pós-graduação em Linguística da Universidade Federal do Rio de Janeiro (UFRJ). E-mail: lorransnm@hotmail.com ORCID: 0000-0003-1409-3597.

1. INTRODUÇÃO

O presente programa experimental visa a investigar estruturas de tópicocomentário como em: “*Aquelas florestas, os arbustos são grandes, por isso eu gostei delas*”, “*Aquelas florestas, os arbustos são grandes, por isso eu gostei deles*”, “*Aquelas florestas, as árvores são grandes, por isso eu gostei delas*” e “*Aquelas florestas, os arbustos são grandes, por isso eu gostei*” no português brasileiro, através de rastreamento ocular. O tópico corresponde a um elemento voltado para o discurso e que não necessariamente concorda com o verbo, majoritariamente localizado na periferia esquerda da sentença, tendo características diferentes em relação às estruturas de sujeito-predicado, quanto a sua sintaxe, prosódia e pragmática (cf. Pontes, 1987). Das construções de tópico no PB, nota-se que, desde o trabalho seminal de Pontes (1987 *apud* Medeiros, 2021), cresce a discussão acerca da presença dos tipos de estrutura de tópico, como se observa nas quatro principais estratégias abaixo:

a) Topicalização

Neste tipo de estrutura, percebe-se um estabelecimento de concordância entre tópico e comentário, em que o elemento topicalizado é um constituinte gerado na posição de argumento interno do núcleo da sentença, subvertendo a ordem SVO original. Ou seja, a estrutura antes orientada para a sentença, por influência da movimentação sintática, é orientada para o discurso.

Isto comprova um vínculo entre o SN de tópico e o predador verbal ou nominal, a ver:

- (1) **Aquele bolo**_i, eu cozinhei ___i.
- (2) **Delicioso**_i, aquele bolo é ___i.

b) Deslocamento à esquerda

Diferente do primeiro tipo de construção de tópico, o deslocamento à esquerda não apresenta categoria vazia, mas, sim, confere uma correferência entre o tópico e um constituinte pronominal no comentário.

- (3) **O Fernando**, eu encontrei **ele** no hospital.
- (4) **O data show**, **ele** depois da queda de energia, **ele** não pega de jeito nenhum

(Exemplos retirados de Silva, 2011:37)

c) Tópico-Sujeito

Aqui, tópico e sujeito constituem um único elemento, de caráter fundido. Pela particularidade de as sentenças deste tipo aparentarem uma estrutura SVO, faz-se a concordância entre tópico e comentário, isto é, tópico/sujeito é um constituinte que estabelece relação argumental com o verbo.

(5) **Essa casa** bate bastante sol.

(6) **A belina** deita o banco, sabe?

(Exemplos retirados de Pontes, 1987:34)

d) Tópico estilo-chinês (Anacoluto/Duplo Sujeito/Tópico pendente)

Em contraposição a outros tipos de estrutura, o tópico de estilo-chinês não é parte constituinte da predicação verbal da sentença, apesar de estabelecer relação semântica com o comentário. Ele parece, então, se caracterizar numa impossibilidade de movimentação sintática (cf. Medeiros, 2021, p. 27-28):

(7) **Filme**, eu gosto mais de comédia.

(Exemplo retirado de Berlinck, Duarte & Oliveira, 2015:82)

(8) **Os museus**, os prédios tudo entregue às baratas não têm um cuidado né?

(Exemplos retirados de Silva, 2006:53)

(9) **Aquelas cidades**_i, os mercados são caros, por isso eu **as**_i detestei.

Outra questão importante de se observar no tópico de estilo-chinês é que ele leva essa nomenclatura justamente pela sua semelhança com a construção do tópico na língua chinesa, voltada ao discurso, como analisa Medeiros (2021, p. 28). Além desse nome, a autora organiza outras classificações que se fazem a esse tipo de estrutura, tais como *Tópico pendente* (cf. Berlinck, Duarte & Oliveira, 2015), *Anacoluto*, pela tradição gramatical (cf. Bechara, 2010) ou *Duplo sujeito* (cf. Li & Thompson, 1976).

Quanto ao contexto pelo qual o tópico se insere, Van Dijk (1982, *apud* Pontes, 1985, p. 54) declara tal constituinte como informação velha, “já sabida”, levando em consideração que o conteúdo semântico seria já conhecido pelos interlocutores em um diálogo. Em comparação com outro constituinte que desempenha função discursiva na periferia esquerda da sentença, o foco, observa-se nas sentenças abaixo, respectivamente

de tópico e foco, como o elemento tópico se diferencia do foco no que tange ao *status* informacional:

(10) A Maria, [FLex o João comprou flores para ela].

(11) Para a Maria [[FLex o João comprou flores, não para a Joana].

(Exemplo retirado de Kato & Mioto, 2015:38)

Em (10), a sentença indica o constituinte [A Maria] como elemento topicalizado, informação antiga, contrastando com a sentença (11) de foco, por exemplo, em que o elemento focalizado é uma informação nova. Com a adição de “Para a” e a substituição de “para ela” por “não para a Joana” na sentença, enfatiza-se que as flores foram compradas para ninguém mais do que Maria e qualquer sugestão contrária, qualquer possível ambiguidade na sentença, é dissipada.

Segundo a classificação tipológica das línguas naturais proposta por Li e Thompson (1976), dividem-se as línguas em quatro tipos, de acordo com a sua orientação, para o discurso e/ou para a sentença. São estas:

- i. Línguas de sujeito proeminente, como as línguas da família indo-europeia;
- ii. Línguas de tópico proeminente, como o mandarim;
- iii. Línguas mistas, isto é, de tópico e sujeito proeminentes, como o japonês e coreano,
- iv. Línguas sincréticas, isto é, sem proeminência de tópico e sujeito, como o tagalog.

Em relação à caracterização do PB como língua SVO, estudiosos como Duarte (1996) e Kenedy (2002, 2014) atribuem à língua a estrutura de sujeito-predicado como a ordenação padrão. Nessa concepção, as estruturas de tópico-comentário são consideradas como desvios gramaticais, tendo em vista que o PB, como língua voltada à sentença, deve ter como ordem predominante das sentenças o Sujeito-Predicado. No entanto, ainda há controvérsias sobre qual seria a classificação mais adequada para o PB. Estudos vêm demonstrando o alto grau de aceitabilidade e produtividade do tópico-comentário pelos nativos da língua, a fim de trazer à discussão a hipótese de uma língua com proeminência tanto de sujeito quanto de tópico. Pesquisas têm sido realizadas desde o trabalho seminal de Pontes (1987), tanto no âmbito descritivo (cf. Duarte, 1996; Negrão, 1999; Galves, 2001; Orsini, 2003; Kato, 2006), quanto no âmbito experimental (cf. Kenedy, 2002, 2014; Medeiros, 2021, entre outros), com o objetivo de se verificar qual seria o *status* tipológico do PB, seguindo tal caracterização.

Li e Thompson (1976) apontam uma série de características que podem ser utilizadas para a verificação da tipologia de uma língua. Dentre tais características, os autores consideram que o controle da correferência dentro de uma estrutura de tópico-comentário pode funcionar como uma espécie de diagnóstico para se verificar a tipologia da língua, em relação à sua proeminência, para o discurso ou para a sentença (p. 469). Isto é, em línguas de tópico proeminente como o chinês, numa estrutura de tópico-comentário como a que se encontra abaixo, a atribuição do controle da correferência ocorre por meio do elemento tópico:

(12) Nei kuai tian daozi zhangde hen da,
aquele pedaço terra arroz crescer muito grande

suoyi _____ hen zhiqian.
então _____ muito valioso

"Aquele pedaço de terra (tópico), o arroz cresce muito grande, então **ela** (a terra) é muito valiosa."

(Exemplo retirado e adaptado de Li & Thompson, 1976:469)

Medeiros (2021), ao analisar a sentença (12) em mandarim, verifica que o controle da correferência se dá pelo tópico, como apontam Li & Thompson (1976), tendo em vista que o constituinte pronominal [ela] estabelece concordância com o tópico [Aquele pedaço de terra] e não o sujeito [arroz].

Kenedy & Mota (2012) ao investigarem tal questão, consideraram que o PB apresenta características de uma língua de sujeito proeminente, a partir dos resultados dos experimentos realizados no âmbito de sua investigação. Sua pesquisa apresentou um experimento de julgamento imediato de gramaticalidade de tópico-sujeito como, por exemplo, "Aquele secretária de vermelho, o diretor disse que demitiu", em *design 2x2*, a fim de testar a hipótese de que o PB é língua com proeminência de tópico, como apontam Pontes (1987) e Kato (2006). Para a tarefa do experimento, os autores investigaram as seguintes condições experimentais:

(a) Referente em posição de sujeito, retomada com anáfora nula (Szero)

“[Aquele secretária de vermelho]_i disse que o diretor demitiu Ø_i”

(b) Referente em posição de sujeito, retomada com anáfora pronominal (Spro)

“[Aquela secretária de vermelho]_i disse que o diretor demitiu ela_i”

(c) Referente em posição de tópico, retomada com anáfora nula (Tzero)

“[Aquela secretária de vermelho]_i, o diretor disse que demitiu Ø_i”

(d) Referente em posição de tópico, retomada com anáfora pronominal (Tpro)

“[Aquela secretária de vermelho]_i, o diretor disse que demitiu ela_i”

O participante lê uma dessas sentenças acima na tela do computador e, em sequência, duas imagens de desenhos humanos são expostas, de modo que a primeira faz referência ao conteúdo do tópico e a segunda ao conteúdo do sujeito. Após a leitura, a pergunta a ser respondida é se a frase lida é considerada aceitável ou inaceitável. Encontrou, em seus resultados, baixo índice de aceitação dos referentes em posição de tópico e custo de processamento do elemento tópico maior em comparação ao sujeito. Os autores consideram, então, que o comportamento dos participantes de sua pesquisa não foi altamente favorável para entender a língua como orientada para o discurso, mas sim para o sujeito, apontando a necessidade de que mais dados linguísticos sejam recolhidos a fim de haver conclusão mais assertiva acerca dessa discussão.

Em contraste, dados coletados na pesquisa de Medeiros (2021) comprovam que o tópico, presente em línguas como o chinês, uma língua orientada para o discurso, são aceitas naturalmente por falantes nativos de PB, tanto quanto o sujeito. A autora, então, defende que o PB seria uma língua mista. Medeiros (2021) considera que o estudo psicolinguístico de Kenedy & Mota (2012) sobre o controle da correferência no PB não levou em consideração um ponto chave na investigação, que é a observação do fenômeno dentro de uma estrutura onde ambos os elementos, tópico e sujeito estejam presentes, para que, assim, possa ser verificada a questão do controle correferência na sentença, de fato. De acordo com Medeiros (2021), isso pode ser problemático e pode ter gerado um resultado inconclusivo sobre o caso do PB. O presente estudo tem o objetivo de verificar tal questão, trazendo mais dados empíricos, oriundos de um teste de rastreamento ocular, com o objetivo de contribuir com a investigação do fenômeno da correferência em estruturas de tópico-comentário do tipo Duplo sujeito (ou tópico estilo-chinês), seguindo a hipótese de que o PB seria uma língua de tópico ou, no mínimo, mista.

O programa experimental conduzido terá como foco a realização de experimentos na área da psicolinguística, nos quais serão verificados o processamento e a interpretação das estruturas de tópico-comentário no PB, por meio de testes de leitura das estruturas

utilizando-se rastreamento ocular e um teste de *cross-modal priming*, com *input* auditivo. O presente estudo relata os procedimentos e resultados obtidos em experimento de rastreamento ocular, já conduzido.

Para trazer maiores contribuições tanto com a literatura em psicolinguística e com a literatura teórica sobre o tópico no PB, pretende-se verificar o comportamento dos participantes durante a tarefa utilizando-se *input* oral e visual. Pretende-se preencher a lacuna existente na literatura, a fim de trazer mais dados oriundos de experimentação psicolinguística, para apoiar ou refutar tanto Kenedy & Mota (2012), quanto Medeiros (2021) e fazer avançar os estudos sobre o PB e sua gramática através de metodologia científica.

2. METODOLOGIA

Para a investigação da questão proposta conduzimos um teste de rastreamento ocular durante a leitura de sentenças de tópico-comentário do tipo estilo-chinês. O objetivo deste experimento foi investigar como ocorre o processamento *on-line* dessas estruturas por falantes nativos de PB. Além disso, pretendeu-se verificar qual elemento controlaria a correferência dentro das estruturas em questão: se é o tópico ou o sujeito, aferindo assim mais diretamente a tese de Li & Thompson (1976) sobre este fenômeno. As frases foram lidas pelos participantes e após a leitura, os participantes responderam a uma pergunta interpretativa. Dessa forma, os movimentos sacádicos progressivos e regressivos, as fixações nas áreas críticas, bem como as interpretações e os tempos de decisão foram computados. A seguir, reporta-se a metodologia, os procedimentos realizados, bem como os achados do teste de rastreamento ocular.

2.1 HIPÓTESES E PREVISÕES

A hipótese adotada nesta pesquisa é a de que o PB seria uma língua mista, portanto, também apresentaria características de uma língua de tópico proeminente, como é o caso do chinês, em relação ao controle da correferência pelo tópico. Sendo assim, os moldes da tradição gramatical para a tipologia do PB não se sustentariam, tendo em vista a não proeminência de apenas sujeito na língua. Conforme Li & Thompson (1976), em uma língua de tópico proeminente, o que controla a correferência numa estrutura de tópico-comentário seria o elemento tópico e não o sujeito (cf. Li & Thompson, 1976). Os autores consideram que em uma língua de tópico proeminente a correferência se dá

apenas ao elemento tópico, paralelamente a uma língua de sujeito proeminente, a qual o controle da correferência se dá pelo sujeito da estrutura. Se o PB, uma língua tradicionalmente classificada como língua de proeminência ao sujeito, é capaz de produzir sentenças em estrutura de tópico e elas são consideradas gramaticais por seus falantes, então ela entraria em contraste com a afirmativa de Li & Thompson, a não ser que o *status* tipológico da língua entre em discussão.

Diante disto, de acordo com a hipótese entretida, espera-se que os participantes não apresentem dificuldades na leitura das sentenças de tópico-comentário, já que o PB seria uma língua mista. Em relação ao fenômeno-alvo deste estudo, espera-se que o tópico desempenhe um papel importante durante a correferência das estruturas, assim como ocorre em línguas de tópico proeminente.

2.2 O *DESIGN* EXPERIMENTAL

O experimento apresenta um *design* experimental 4, onde manipulou-se o fator *Tipo de referente*, como variável independente, com 4 níveis, a saber: **Tópico x Sujeito x Ambíguo x Nulo**, gerando-se assim, quatro condições experimentais: **TOP, SUJ, AMB** e **NUL**, conforme exemplos abaixo:

Tabela 1: Exemplos das condições experimentais.

Aquelas florestas, os arbustos são grandes, por isso eu não gosto delas . (TOP)
Aquelas florestas, os arbustos são grandes, por isso eu não gosto deles . (SUJ)
Aquelas florestas, as árvores são grandes, por isso eu não gosto delas . (AMB)
Aquelas florestas, os arbustos são grandes, por isso eu não gosto _____. (NUL)

Fonte: as autoras (2024)

As variáveis dependentes foram os tempos de fixação ocular nas regiões críticas das sentenças, o padrão de movimentação sacádica (movimentos progressivos e regressivos), como medidas *on-line*, e a preferência de correferência pelos elementos tópico e sujeito nas estruturas de tópico, como medida *off-line*, bem como os tempos desta escolha (*off-line* cronométrico).

Espera-se que os participantes não apresentem dificuldades durante a leitura das condições *TOP*(Tópico) e *SUJ*(Sujeito), em que o pronome anafórico esteja sendo correferente ao tópico “Aquelas florestas” e ao sujeito “os arbustos”, respectivamente, tanto no processamento *on-line* das estruturas, quanto na avaliação das questões interpretativas, já que a hipótese entretida neste experimento é a de que o PB seria uma língua mista. A condição *AMB*(ambígua) será utilizada como espécie de sonda para verificar se a preferência dos participantes será o tópico ou sujeito numa situação em que ambos os elementos estão disponíveis. A condição *NUL*(Nulo) verificará qual o elemento preferido para ser o referente numa situação em que há uma lacuna sem ser preenchida. Dessa maneira, as condições *Ambígua* e *Nulo* nos permitirão verificar qual elemento controlaria a correferência nas estruturas de tópico-comentário testadas, mais diretamente.

2.3 PARTICIPANTES

Participaram deste experimento 40 alunos do curso de Letras da Universidade Federal do Rio de Janeiro – UFRJ, com idades entre 18 e 33 anos, com visão normal ou corrigida. Cada participante recebeu 2 horas de créditos complementares (AACC) pela sua participação.

2.4 MATERIAIS

Para compor o material, foram confeccionados 16 conjuntos experimentais, além de 32 frases distratoras. Antes de cada sentença, uma imagem foi apresentada a fim de oferecer um contexto prévio para situar o participante, já que estruturas de tópico-comentário são consideradas como um fenômeno da língua oral. No entanto, tais imagens não foram utilizadas como variáveis a serem analisadas, apenas como uma espécie de contexto prévio para licenciar o tópico, já que tal construção tem forte relação com o discurso.

Nas frases experimentais controlaram-se algumas variáveis importantes, tais como: animacidade, gênero, número de sílabas, uso de preposições, além do tamanho das sentenças. Dessa forma, trabalhamos com tópicos e sujeitos inanimados; metade dos tópicos e dos sujeitos sendo do gênero feminino e metade sendo do gênero masculino; os adjetivos eram dissílabos; SN1 e SN2 compostos por artigo determinante + substantivo trissílabo, e metade do conjunto apresentando preposição "de" + pronome "eles"/"elas" e

metade apresentando os pronomes clíticos "os"/"as". Abaixo, há dois conjuntos experimentais que ilustram tais controles, a servir de exemplo:

Quadro 1: Exemplos dos materiais experimentais (conjunto 3).



Fonte: as autoras (2024)

- Aquelas pinturas, os desenhos são legais, por isso eu gosto delas.
- Aquelas pinturas, os desenhos são legais, por isso eu gosto deles.
- Aquelas pinturas, as texturas são legais, por isso eu gosto delas.
- Aquelas pinturas, os desenhos são legais, por isso eu gosto _____.

Questão: Do que eu gostei?

a) Pintura b) Desenho/Texturas

Quadro 2: Exemplos dos materiais experimentais (conjunto 15).



Fonte: as autoras (2024)

- Aqueles óculos, as armações são fracas, por isso eu os quebrei.
- Aqueles óculos, as armações são fracas, por isso eu as quebrei.
- Aqueles óculos, os modelos são fracos, por isso eu os quebrei.
- Aqueles óculos, as armações são fracas, por isso eu __ quebrei.

Questão: O que eu quebrei?

a) Óculos b) Armações/Modelos

Os materiais experimentais foram distribuídos de acordo a configuração quadrado latino, de forma que todos os participantes fossem expostos a todas as condições experimentais, mas garantindo que não fossem expostos a sentenças com o mesmo conteúdo lexical. As frases foram randomizadas automaticamente. Cada participante leu um total de 48 sentenças, sendo 16 frases experimentais e 32 frases distratoras. As frases apareciam na tela do computador, com fundo branco, fonte *Monaco* 24, na cor preta, aparecendo em uma única linha, da esquerda para a direita. As questões interpretativas eram apresentadas em uma linha e as duas opções de resposta logo abaixo, com a opção SIM à esquerda e a opção NÃO, à direita.

2.5 PROCEDIMENTOS

Conforme mencionado, o experimento utilizou a técnica de rastreamento ocular (*Eye-tracking*). O rastreador usado para captura dos movimentos oculares e registro dos padrões de leitura foi um *EyeLink* 1000 (SR Research), com uma câmera de alta precisão com acurácia de 1000Hz, configurada para gravação monocular. A tela onde os participantes viam os estímulos tinha 32 polegadas, com resolução de 1920x1080 *pixels*. Como o aparelho utilizado é de alta sensibilidade, utilizou-se um aparato para cabeça e testa do participante, garantindo que ele não movimentasse a cabeça durante a tarefa, enquanto seus movimentos oculares eram registrados. O programa usado para apresentação e registro dos estímulos foi o *Experiment Builder* (SR Research) e a análise foi feita utilizando-se o programa *Data Viewer* (SR Research), ambos nativos do sistema *Eyelink*.

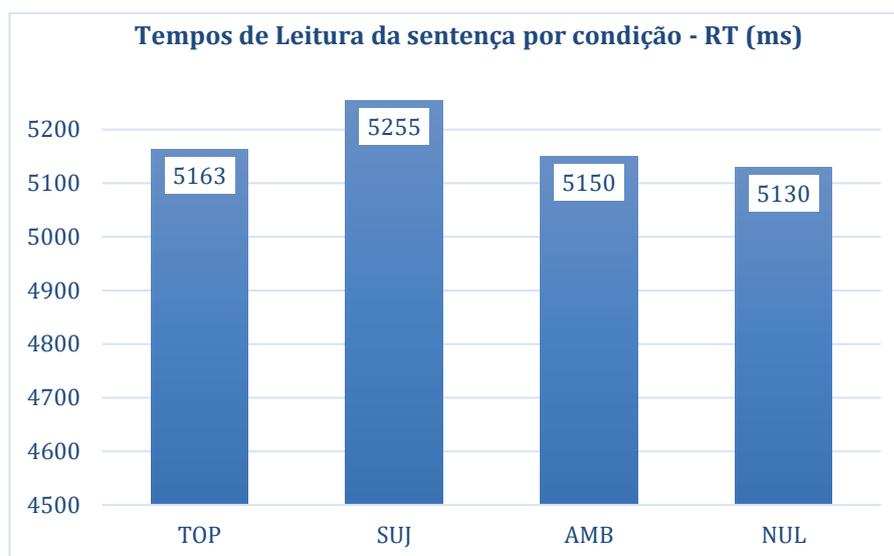
O experimento foi aplicado no Laboratório de Psicolinguística Experimental – LAPEX, na Faculdade de Letras da UFRJ. A tarefa realizada pelos participantes era

simples: primeiro, os participantes viam uma imagem na tela que tinha relação com a frase a ser lida (contexto prévio). Esta imagem ficava na tela durante 3 segundos. Em seguida, a imagem sumia automaticamente e aparecia a frase. Os participantes faziam a leitura da frase silenciosamente e após a leitura, apertavam a tecla “Espaço” no teclado, para ler a pergunta interpretativa. Para responder a opção da esquerda, apertavam a tecla “S”, marcada na cor verde, e para responder a opção da direita, apertavam a tecla “L”, marcada na cor amarela. Antes do experimento ter início, de fato, o experimentador dava todas as instruções necessárias aos participantes para a realização da tarefa e fazia o procedimento de calibragem da pupila dos participantes no rastreador, para correta captura dos seus movimentos oculares. Após todos os procedimentos pré-testes, o experimento era iniciado.

3. RESULTADOS

Após a captura dos dados, realizou-se a etapa de preparo dos materiais para a extração das medidas, que consistia em separar as frases em regiões de interesse (ROIs) específicas, para visualização e extração das métricas de leitura. Selecionou-se como região de interesse os elementos tópico e sujeito das estruturas, além dos pronomes que faziam referência aos elementos em questão, sendo, em certas condições, ambíguos e nulos. As medidas que serão reportadas a seguir referem-se a cada uma dessas regiões. Em relação a análise, quando se trata da medida “*Tempo total de leitura da frase*”, realizou-se uma ANOVA com *design 1 within subjects*, tendo o fator “*Condição*” como alvo da análise, com quatro níveis: *Tópico x Sujeito x Ambíguo x Nulo*. Para a análise da medida “*Primeiro passe de leitura*” e “*Segundo passe de leitura*”, do tópico e do sujeito, realizou-se uma ANOVA com *design 2 within subjects*, cruzando-se os fatores “*Condição*”, com quatro níveis (*Tópico x Sujeito x Ambíguo x Nulo*) e “*Referente*”, com dois níveis (*Tópico x Sujeito*). Já na análise do “*Primeiro passe de leitura*” e “*Segundo passe de leitura*”, dos pronomes, realizou-se ANOVA com *design 1 within subjects*, tendo o fator “*Condição*” três níveis: *Tópico x Sujeito x Ambíguo*. A condição *Nulo* não entrou nesta análise, pois não houve dado a ser observado, visto que em tal condição o pronome era inexistente. Tais dados serão reportados a seguir.

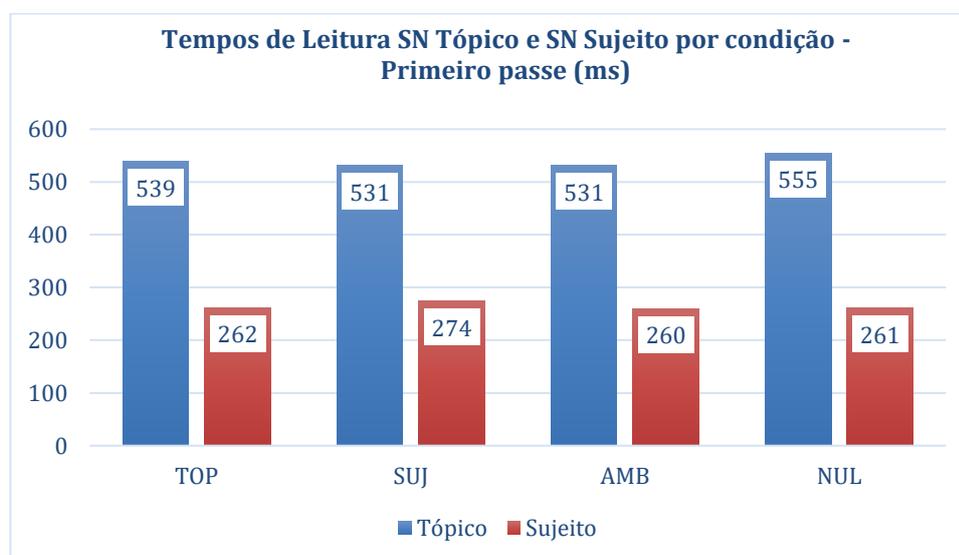
O gráfico abaixo mostra o tempo total de leitura das frases nas quatro condições experimentais investigadas:

Gráfico 1: Tempo de leitura das sentenças por condição

Fonte: as autoras (2024)

O resultado da ANOVA não indicou efeito principal de Condição $F(3,477) = 1,39$ $p < 0,245656$, não sendo observada, ainda, diferença nos tempos de leitura durante a comparação por pares. Isto é, comparando-se os tempos totais de leitura nas quatro condições testadas, não houve diferença estatística significativa, indicando que todas as sentenças foram lidas da mesma maneira.

O gráfico 2 mostra os tempos do primeiro passe da leitura (sem captura de regressões) apenas dos elementos tópico e sujeito, nas condições testadas.

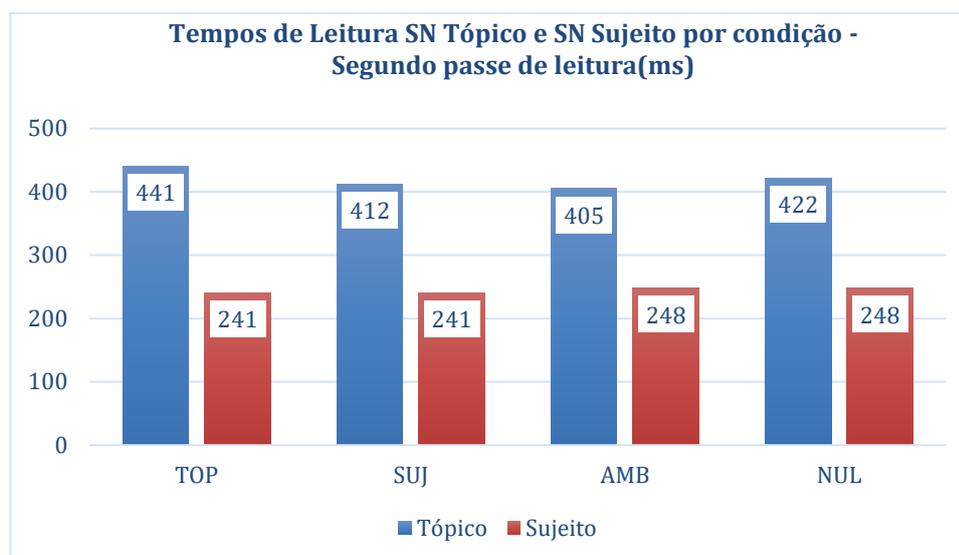
Gráfico 2: Primeiro passe de leitura nas regiões de interesse por condição

Fonte: as autoras (2024)

O resultado da ANOVA não indicou efeito principal de Condição $F(3,477) = 1,39$ $p < 0,245656$, mas indicou efeito principal de Referente $F(1,159) = 2522$ $p < 0,000001$ altamente significativo. Houve interação entre os fatores Condição*Referente $F(3,477) = 3,35$ $p < 0,018907$. Comparando-se os pares, os testes-t indicaram diferença estatística altamente significativa em todas as comparações relevantes: [TOP_Tópico]vs[TOP_Sujeito] $t(159)=31,97$ $p < 0,0001$; [SUJ_Tópico]vs[SUJ_Sujeito] $t(159)=26,97$ $p < 0,0001$; [AMB_Tópico]vs[AMB_Sujeito] $t(159)=29,34$ $p < 0,0001$ e [NUL_Tópico]vs[NUL_Sujeito] $t(159)=33,13$ $p < 0,0001$. Tais dados indicam que independentemente da condição, o referente "tópico" é lido mais custosamente durante o primeiro passe do que o referente "sujeito" em todas as estruturas.

O gráfico 3 mostra os tempos do segundo passe da leitura (movimentos sacádicos regressivos, indicando releitura da região) dos elementos tópico e sujeito, nas sentenças em questão.

Gráfico 3: Segundo passe de leitura (releitura) nas regiões de interesse por condição



Fonte: as autoras (2024)

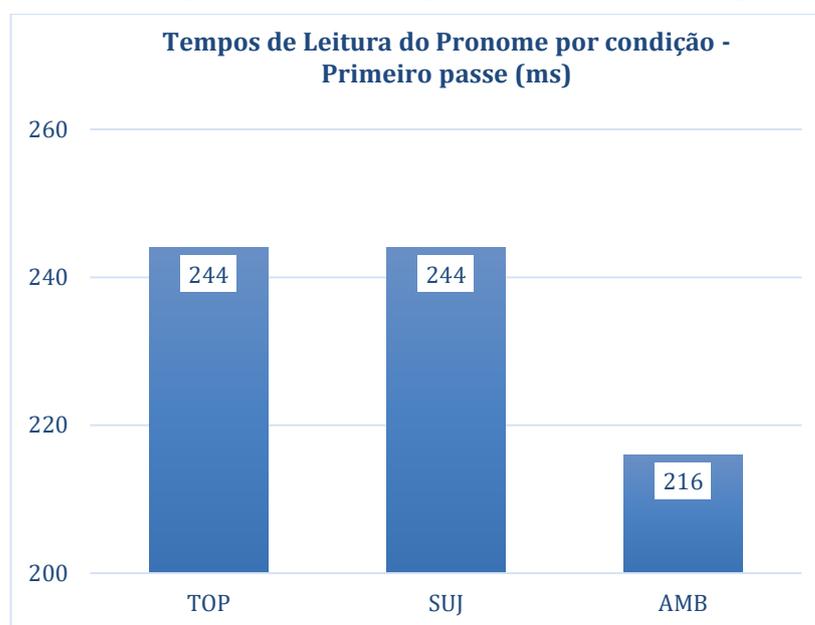
A ANOVA indicou efeito principal de Condição $F(3,477) = 3,50$ $p < 0,015517$ e indicou efeito principal de Referente altamente significativo $F(1,159) = 2598$ $p < 0,000001$, havendo interação entre os fatores Condição*Referente $F(1,159) = 2598$ $p < 0,000001$. Comparando-se os pares, os testes-t também indicaram diferença estatística altamente significativa em todas as comparações: [TOP_Tópico]vs[TOP_Sujeito] $t(159)=26,01$ $p < 0,0001$; [SUJ_Tópico]vs[SUJ_Sujeito] $t(159)=24,13$ $p < 0,0001$; [AMB_Tópico]vs[AMB_Sujeito] $t(159)=23,41$ $p < 0,0001$,

[NUL_Tópico]vs[NUL_Sujeito] $t(159)=23,10$ $p < 0,0001$. Tal padrão foi o mesmo observado na medida de primeiro passe de leitura, confirmando que o referente "tópico" é lido mais custosamente do que o referente "sujeito", em todas as estruturas.

Além dos referentes tópico e sujeito, também se observaram os tempos de leitura e releitura dos pronomes, pois esses elementos faziam a retomada tanto do tópico, quanto do sujeito, a depender da condição.

O gráfico 4 mostra os tempos do primeiro passe da leitura dos pronomes nas frases lidas. É importante salientar que os gráficos referentes ao pronome têm apenas 3 condições em comparação com a medida da frase total, pois há uma condição onde não há ocorrência de pronome (NULO). Vejamos o resultado a seguir:

Gráfico 4: Primeiro passe de leitura nas regiões de interesse (Pronome) por condição



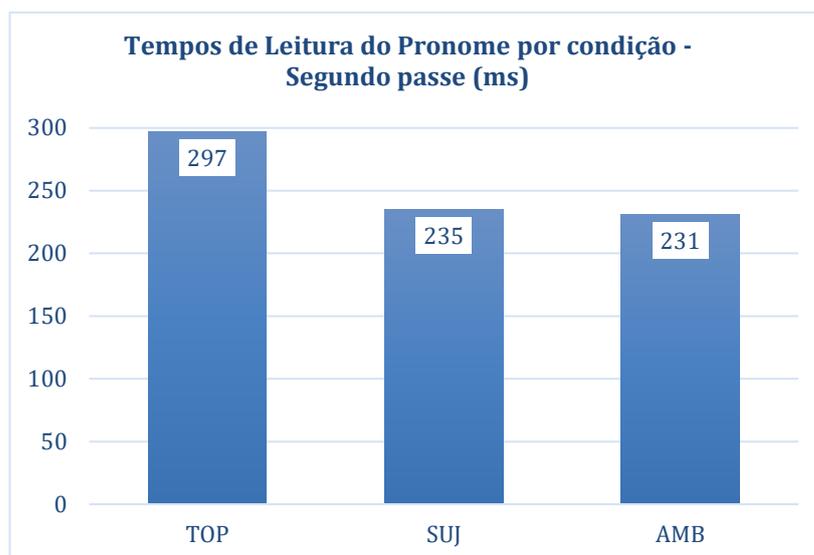
Fonte: as autoras (2024)

A ANOVA indicou efeito principal de Condição $F(2,318) = 29,1$ $p < 0,000001$ altamente significativo. Comparando-se os pares, os testes-t indicaram diferença estatística significativa nas comparações: [TOP]vs[AMB] $t(159)=7,11$ $p < 0,0001$ e [SUJ]vs[AMB] $t(159)=6,81$ $p < 0,0001$. Na comparação [TOP]vs[SUJ] não houve diferença estatística significativa $t(159)=0,01$ $p < 0,9968$. Tal resultado indica que, durante o primeiro passe de leitura do pronome, a condição "Ambígua" teve menor tempo de processamento em comparação às demais condições. Isto pode indicar que, como qualquer uma das opções disponíveis, tópico ou sujeito, poderia ser possível, o *parser*

não teve “trabalho” durante a leitura desta estrutura, o que ocasionou os baixos tempos de leitura observados.

No gráfico 5 é possível observar os tempos do segundo passe da leitura dos pronomes:

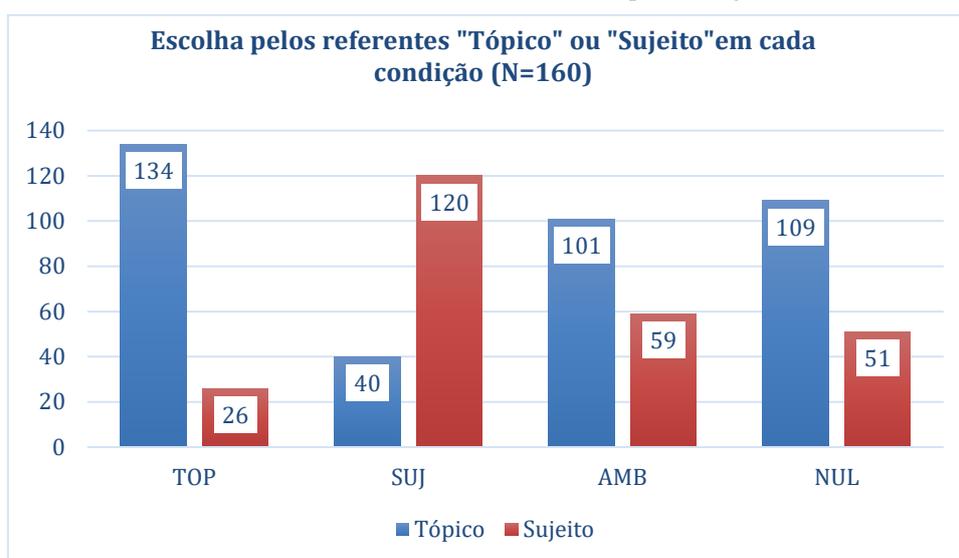
Gráfico 5: Segundo passe de leitura (releitura) nas regiões de interesse (Pronome) por condição



Fonte: as autoras (2024)

A ANOVA indicou efeito principal de Condição $F(2,318) = 775$ $p < 0,000001$ altamente significativo. Na comparação por pares, os testes-t indicaram diferença estatística significativa em todas as comparações: [TOP]vs[SUJ] $t(159)=33,24$ $p < 0,0001$; [TOP]vs[AMB] $t(159)=35,35$ $p < 0,0001$ e [SUJ]vs[AMB] $t(159)=2,14$ $p < 0,0335$. Tal resultado demonstra que houve mais regressões na região do pronome na condição em que o referente era o elemento topicalizado, sugerindo que a preferência de correferência se deu para o tópico, ainda no processamento *on-line*.

Na análise dos dados da medida *off-line*, os índices de respostas obtidos mostram que, na escolha do referente, os participantes preferiram escolher o elemento "tópico" como o referente da sentença quando não havia uma opção clara de escolha (condições Ambígua e Nula), conforme pode ser observado no gráfico 6.

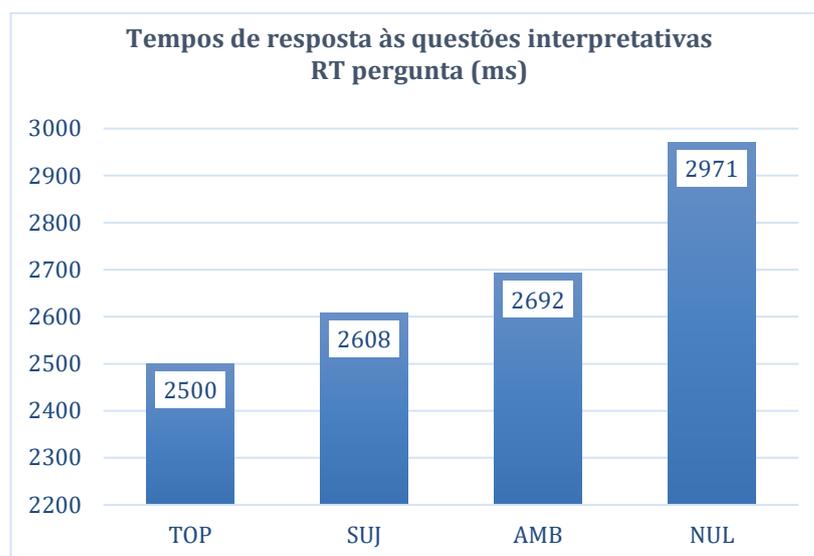
Gráfico 6: Índice de escolha dos referentes por condição

Fonte: as autoras (2024)

O experimento contou com 640 questões interpretativas, feitas pelos 160 participantes da pesquisa. Destas, 384 foram respondidas em favor do tópico e 256 foram respondidas em favor do sujeito. O teste Chi-quadrado indica que esta diferença geral, em favor da construção de tópico, é altamente significativa ($X^2 = 51,2$, $p = 0,0001^{****}$).

Conforme pode ser verificado no gráfico acima, os participantes escolheram fazer a relação pronome-referente ao elemento sujeito da estrutura apenas quando era explícita tal relação (como ocorre na condição SUJ). Isto é, apenas na condição SUJ, os participantes escolheram fazer a relação pronome-referente significativamente **ao elemento sujeito da estrutura** ($X^2 = 80$, $p = 0,0001^{****}$). Nas demais condições, tanto quando a relação era explícita para o tópico ($X^2 = 145,8$, $p = 0,0001^{****}$), quanto quando não havia pronome a ser relacionado ($X^2 = 42,05$, $p = 0,0001^{****}$) ou quando havia uma condição ambígua ($X^2 = 22,05$, $p = 0,0001^{****}$), a escolha se deu em vias de tópico, conforme esperado.

No gráfico 7, podemos verificar os tempos de resposta às questões interpretativas. Tal gráfico demonstra que os participantes demoraram mais para escolher o referente das sentenças na condição "Nulo".

Gráfico 7: Tempos de escolha dos referentes pelos participantes, por condição

Fonte: as autoras (2024)

A ANOVA indicou efeito principal de Condição $F(3,477) = 11,9$ $p < 0,000001$ altamente significativo. Na comparação por pares, os testes-t indicaram diferença estatística significativa em todas as comparações: [TOP]vs[NUL] $t(159)=6,93$ $p < 0,0001$; [SUJ]vs[NUL] $t(159)=4,96$ $p < 0,0001$; [AMB]vs[NUL] $t(159)=3,22$ $p < 0,0016$; [TOP]vs[AMB] $t(159)=2,16$ $p < 0,0326$. Tal resultado indica que, quando não há uma opção explícita de correferência, nem uma opção ambígua, os participantes demoraram mais para decidir sobre quem seria o referente, o que seria totalmente condizente com a situação, considerando que as duas posições, tópico e sujeito, estariam “disponíveis”.

A seguir, discutiremos os achados obtidos relacionando-os com estudos anteriores da literatura e com a hipótese central proposta nesta investigação.

4. DISCUSSÃO

Os achados deste experimento de rastreamento ocular trazem evidências contra a hipótese de que o PB seria uma língua com proeminência de sujeito, orientada para a sentença. Os resultados obtidos dos participantes do teste experimental se dividem entre dados de processamento *on-line* e *off-line*. Resultados *on-line* são aqueles que verificam os tempos de leitura dos sintagmas, a fim de comparar os custos de processamento e o grau de gramaticalidade/aceitabilidade das sentenças. Na medida do tempo de leitura da sentença em cada condição dos conjuntos, não houve diferença estatística significativa,

indicando que todas as condições foram lidas da mesma maneira. Já durante o primeiro e segundo passe de leitura do SN Tópico e SN Sujeito, verifica-se que, assim como Medeiros (2021) aponta, há maior custo de processamento do referente tópico do que o referente sujeito em quaisquer condições. Esse custo, no entanto, muda de figura quanto às leituras do pronome em cada condição: no primeiro passe, a condição “Ambígua” apresenta menor tempo de processamento em comparação com o tópico e o sujeito, tendo em vista que não houve diferença na comparação entre tais constituintes. Já durante o segundo passe de leitura ocorrem mais regressões na região do pronome na condição em que o referente era o elemento topicalizado. Tal experimento demonstra que, apesar de o tempo de leitura (*on-line*) apresentar valores mais elevados para o tópico, isso não afeta a preferência pelo tópico (*off-line*) durante a correferência.

Se os dados *on-line* correspondem ao momento exato em que se demora para ler a sentença, os dados *off-line* complementam com o tempo em que se demora para responder à questão que vem a seguir e a resposta dada pelos participantes. Com tais resultados, pode-se entender de forma mais aprofundada qual a preferência de correferente em estruturas de tópico-comentário no PB. Os tempos de resposta às questões interpretativas indicam que os participantes apresentaram maior dificuldade para escolher o referente das sentenças na condição “nulo” e “ambíguo”. Os dados *off-line* obtidos demonstraram que os participantes preferiram o elemento “tópico” como o referente da sentença quando não estavam diante de uma opção explícita. Tais achados parecem indicar que participantes deste estudo apresentam comportamento condizente com o que se espera de falantes de uma língua de tópico proeminente.

Estes resultados aqui obtidos entram em consonância com a hipótese da precursora Eunice Pontes (1987) e os diversos estudos posteriores já citados. Entram, também, em função de debate e complementação aos achados de Kenedy & Mota (2012), onde concluíram que o grau de produtividade do tópico, em especial o tópico-sujeito no PB, não seria suficiente para considerar a língua de proeminência mista. No entanto, como os próprios autores afirmam, é necessário que mais pesquisas e testes sejam feitos para que se comprove a proeminência de tópico na língua. Como estratégia, o programa experimental vigente escolheu analisar o uso do tópico de estilo-chinês justamente por haver relação sintática intrínseca com a principal língua de proeminência de tópico (cf. Li & Thompson, 1976). Os dados coletados dos participantes deste experimento também se diferem da metodologia de Kenedy & Mota, de modo que o comportamento ocular *on-line* de leitura, como passadas progressivas e regressões, também foi levado em

consideração, em conjunto com as respostas *off-line*. Observar tal fenômeno também se prova essencial para compreender o processamento mental desse tipo de sentença, como se discute em seguida.

Os dados deste experimento de rastreamento ocular mostram que, no PB, o controle da correferência se dá pelo tópico em estruturas onde a possibilidade de correferência é explícita para o tópico e também quando há possibilidade de se escolher tanto tópico quanto sujeito como referentes. Dessa forma, os resultados do presente estudo demonstram que sentenças em PB com correferência ambígua e nula apresentam preferência de aposição ao elemento topicalizado.

A aparente diferença nos índices de tempo de decisão e o padrão de fixações oculares observados quando comparamos os elementos tópico e sujeito nas sentenças vão na mesma direção do que foi verificado em trabalhos anteriores na literatura, como o estudo de Medeiros (2021) sobre o processamento de estruturas de tópico-comentário x estruturas de sujeito predicado no PB. A autora também identificou que durante a leitura de frases, as sentenças de tópico-comentário são lidas mais custosamente do que as estruturas de sujeito-predicado. Medeiros (2021) argumenta que tal fenômeno ocorre devido a questões relacionadas à formação das sentenças em questão e não teriam relação com o tópico ser agramatical na língua, muito pelo contrário. Isto é, estruturas de tópico-comentário possuem um elemento posicionado na periferia esquerda da sentença, sendo esta uma posição não-argumental, que não atribui caso nem papel temático ao elemento que nela se encontra. Ao identificar um elemento ocupando a periferia esquerda da sentença, o *parser* precisa realizar algumas estratégias para interpretação deste elemento, o que não ocorre com o sujeito, por exemplo, que ocupa uma posição argumental na estrutura sintática (cf. Frazier & Clifton, 1989). Como estratégia, o *parser* tenta realocar o constituinte na primeira lacuna disponível, o que gera incrementos nos tempos de processamento, quando comparado a uma estrutura de sujeito.

Portanto, em consonância com os estudos de Medeiros (2021), Maia (1997) e Pontes (1987), somando-se aos achados obtidos neste estudo, defende-se a classificação do PB como língua mista. De fato, os participantes só escolheram o sujeito como referente quando havia uma relação explícita entre tal elemento e seu correferente. Isto pode ser mais uma evidência em favor da classificação tipológica proposta.

5. CONSIDERAÇÕES FINAIS

Em trabalho seminal, Li & Thompson (1976) consideram que a atribuição do controle da correferência dentro de uma estrutura de tópico-comentário pode funcionar como diagnóstico para se verificar a tipologia de uma língua, em relação a sua proeminência, para o discurso ou para a sentença. No que diz respeito ao PB, a tradição gramatical costuma considerar somente a estrutura de sujeito-predicado como a ordenação padrão na língua e as estruturas de tópico-comentário podem ser consideradas até como desvios gramaticais.

Kenedy & Mota (2012) ao investigarem a questão, consideraram que o PB apresenta características de uma língua de sujeito proeminente, a partir dos resultados dos experimentos realizados no âmbito de sua investigação. No entanto, dados coletados na pesquisa de Medeiros (2021) sugerem que o tópico, presente em línguas como o chinês, uma língua orientada para o discurso, é aceito naturalmente por falantes nativos de PB, tanto quanto o sujeito, sendo este um forte traço para se considerar o PB como uma língua mista, nos moldes de Pontes (1987).

Esta pesquisa investigou o processamento de construções de tópico-comentário, em especial o tópico de estilo-chinês, por meio de um teste experimental de rastreamento ocular, com vistas a buscar evidências experimentais que sustentem a hipótese de que o PB é uma língua de proeminência de sujeito e tópico, segundo a classificação tipológica de Li & Thompson (1976). Os resultados obtidos no experimento de rastreamento ocular corroboram tal hipótese, ao observar que os participantes preferiram o elemento “tópico” como o referente da sentença quando não estavam diante de uma opção explícita, apresentando comportamento condizente com o que se espera de falantes de uma língua de tópico proeminente.

As previsões formuladas no início deste trabalho puderam ser confirmadas, tendo em vista que os resultados obtidos demonstram que o elemento tópico tem grande relevância ao considerarmos a questão da correferência nas estruturas de tópico-comentário. Estes achados ressaltam a relevância dos resultados obtidos até o momento e antecipam as expectativas para uma próxima investigação, apontando para uma compreensão mais aprofundada do papel do tópico na correferência no PB, considerando-se também um *input* auditivo, através da técnica de *priming*.

Chinese-style topic and the typological status of BP: an eye tracking study

ABSTRACT:

With the aim of offering evidence that supports the hypothesis of Brazilian Portuguese (BP) as a mixed language, that is, a language that, in addition to presenting subject prominence, also presents topic prominence, the present study analyzes the processing of Chinese-style topic by native BP speakers, through eye tracking. This experiment presents 4 design featuring the factor “Type of referent” with 4 levels that generated the following conditions:

- a) Aquelas florestas, os arbustos são grandes, por isso eu gostei delas. (TOP)
- b) Aquelas florestas, os arbustos são grandes, por isso eu gostei deles. (SUJ)
- c) Aquelas florestas, as árvores são grandes, por isso eu gostei delas. (AMB)
- d) Aquelas florestas, os arbustos são grandes, por isso eu gostei _____. (NUL)

A reading test was conducted, where the hypothesis was that, in a topic language, what would control coreference in a topic-comment structure would be the topic element and not the subject (cf. Li & Thompson, 1976). Therefore, it is expected that participants consider the anaphoric pronoun to be co-referent to the topic. Response times to interpretative questions indicated greater processing time in the “null” and “ambiguous” conditions. The results also indicate that, when there is no explicit relationship between the co-referent element and the pronoun, 65% of choices are made based on the topic, and not the subject. Such data corroborate findings by Medeiros (2021) and refute Kenedy & Mota (2012), who evaluated BP as a language of subject prominence, when they investigated the same phenomenon.

KEYWORDS: Psycholinguistics. Topic-commentary. Eye-tracking. BP.

REFERÊNCIAS BIBLIOGRÁFICAS

DUARTE. A topicalização no português europeu: uma análise comparativa. In: DUARTE, I. & LEIRIA, I. (eds.) **Actas do Congresso Internacional sobre o Português**. Lisboa: APL/Colibri, 1996.

FRAZIER, Lyn; CLIFTON JR, Charles. Successive cyclicity in the grammar and the parser. **Language and cognitive processes**, v. 4, n. 2, p. 93-126, 1989.

GALVES, C. **Ensaio sobre as gramáticas do português**. Campinas: Editora da Unicamp, 2001.

KATO, M. **Comparando o português da América com o português de Portugal e com outras línguas**. 2006.

KENEDY, E. **Aspectos estruturais da relativização em português: uma análise baseada no modelo raising**. (Dissertação de Mestrado). RJ: UFRJ, 2002.

_____.; MOTA, C. **Orientações de anáforas nulas e pronominais para sujeitos e tópicos no PB**. *Linguística*, vol. 8, núm.2. RJ: UFRJ, 2012.

LI, C-N.; THOMPSON, S. Subject and topic: a new typology of language. In: LI, C-N (Ed.). **Subject and Topic**. New York: Academic Press, p. 457-489, 1976.

LIU, R; NICOL, J. **Online processing of Anaphora by advanced English learners**. University of Arizona. 2010.

MAIA, M.. The processing of object anaphora in Brazilian Portuguese. **Révue Linguistique de Vincennes** no 26, p.151-172, França, 1997.

MEDEIROS, L.S.N.V. **Processamento de construções de tópico-comentário e sujeito-predicado no português do Brasil: língua orientada para a sentença, para o discurso ou mista?** Tese (Doutorado em Linguística) – Universidade Federal do Rio de Janeiro, 304 f, 2021.

NEGRÃO, E. V. **O português brasileiro: uma língua voltada para o discurso.** (Tese de livre docência). USP: SP, 1999.

ORSINI, M. T. **As construções de tópico no português do Brasil: uma análise sintático-discursiva e prosódica.** (Tese de Doutorado). UFRJ: RJ, 2003

PONTES, E. Sujeito e tópico do discurso. **DELTA: Documentação e Estudos em Linguística Teórica e Aplicada**, v. 1, n. 1/2, 1985.

_____, E. **O tópico no português do Brasil.** Pontes Editores, 1987.